

THE NLSY79

In 1977, a panel of experts convened by the U.S. Department of Labor recommended that the NLS begin a longitudinal study of a panel of young men and women to permit a replication of the analysis of the 1960s cohorts of young men and young women. A new longitudinal study would also assist in the evaluation of the expanded employment and training programs for youths legislated by the 1977 amendments to the Comprehensive Employment and Training Act (CETA).

To these ends, investigators drew a national probability sample of civilian young men and women aged 14 to 21 (as of December 31, 1978), with overrepresentation of blacks, Hispanics, and economically disadvantaged nonblack/non-Hispanics. Funding from the Department of Defense permitted interviews with an additional group of young persons serving in the Armed Forces. The first interview with this sample of civilian and military youth, called the National Longitudinal Survey of Youth 1979 (NLSY79), took place in early 1979. The cohort was re-interviewed annually through 1994; after 1994, the survey schedule was changed to every other year. Data through the 1998 interview are now available. Data from the 2000 interview are scheduled for release in fall 2001; researchers interested in more details should contact NLS User Services at usersvc@postoffice.chrr.ohio-state.edu.

Since 1986, detailed information on the development of children born to women in the NLSY79 cohort has supplemented the data on mothers and children collected during the regular youth surveys. Chapter 4 describes the biennial surveys of this group, referred to as the Children of the NLSY79.

Chapter organization

This chapter supplies an overview of the main NLSY79 cohort and the types of data present within this data set. It provides information about the main and supplemental NLSY79 samples and includes a discussion of multiple-respondent households in the NLSY79, an important feature of the original sampling design. This chapter also gives information on interviews with the NLSY79 men and women, as well as information relevant to assessing how attrition has affected the sample sizes over time. Major data elements for the NLSY79 are then highlighted. The chapter

concludes with a brief summary of the data files and documentation available for the NLSY79.

The NLSY79

The NLSY79 is a nationally representative sample of 12,686 young men and women who were 14 to 22 years of age when first surveyed in 1979. During the years since that first interview, these young people typically have finished their schooling, moved out of their parents' homes, made decisions on continuing education and training, entered the labor market, served in the military, married, and started families of their own. Data collected from the NLSY79 respondents chronicle these changes, providing researchers with a unique opportunity to study in detail the life course experiences of a large group of adults representative of all men and women born in the late 1950s and early 1960s and living in the United States when the survey began.

The NLSY79 samples

The NLSY79 sampling design enables researchers to analyze the disparate life course experiences of groups such as women, Hispanics, blacks, and the economically disadvantaged. Three subsamples make up the NLSY79:

- A cross-sectional sample of 6,111 youths designed to be representative of noninstitutionalized civilian youths living in the United States in 1979 and born between January 1, 1957, and December 31, 1964 (aged 14 to 21 as of December 31, 1978)
- A supplemental sample of 5,295 youths designed to oversample civilian Hispanic, black, and economically disadvantaged nonblack/non-Hispanic youths living in the United States during 1979 and born between January 1, 1957, and December 31, 1964
- A military sample of 1,280 youths designed to represent the population born between January 1, 1957, and December 31, 1961 (aged 17 to 21 as of December 31, 1978), and enlisted in one of the four branches of the active military forces as of September 30, 1978

All members of the cross-sectional sample and the Hispanic and black supplemental samples have been eligible for interview during each round of the NLSY79. Funding con-

straints limited the numbers of military sample members interviewed after 1984 and of economically disadvantaged nonblack/non-Hispanic supplemental sample members interviewed after 1990. (See table 3.1.)

Multiple respondent households

The NLSY79 sample design selected as respondents all eligible individuals aged 14 to 21 who resided in a surveyed household at the end of 1978. As a result, the 11,406 civil-

ian respondents interviewed in 1979 originated from 7,490 unique households; 2,862 households included more than one NLSY79 respondent.

At the time the surveys began, the primary types of relationships between respondents living in multiple-respondent households were those of spouse and biological, step-, or adopted sibling. Table 3.2 lists the number of cohort members living in multiple respondent households during the initial survey.

Table 3.1. Distribution of NLSY79 respondents by sample type, race/ethnicity, and sex

Sample type	1979	1984	1990	1994	1998
Total sample	12,686	12,069	10,436	8,891	8,399
Cross-sectional sample	6,111	5,814	5,498	5,457	5,159
Males	3,003	2,839	2,664	2,648	2,459
Nonblack/non-Hispanic	2,439	2,303	2,157	2,150	2,007
Black	346	329	318	310	289
Hispanic	218	207	189	188	163
Females	3,108	2,975	2,834	2,809	2,700
Nonblack/non-Hispanic	2,477	2,365	2,271	2,243	2,161
Black	405	393	365	363	346
Hispanic	226	217	198	203	193
Supplemental sample	5,295	5,040	4,755	3,256	3,065
Males	2,576	2,442	2,280	1,599	1,480
Economically disadvantaged nonblack/non-Hispanic	742	699	664	1—	1—
Black	1,105	1,055	979	973	883
Hispanic	729	688	637	626	597
Females	2,719	2,598	2,475	1,657	1,585
Economically disadvantaged nonblack/non-Hispanic	901	851	819	1—	1—
Black	1,067	1,034	984	987	951
Hispanic	751	713	672	670	634
Military sample	1,280	1,215	² 183	178	175
Males	824	774	168	164	161
Nonblack/non-Hispanic	609	575	70	68	66
Black	162	151	68	62	63
Hispanic	53	48	30	34	32
Females	456	441	15	14	14
Nonblack/non-Hispanic	342	331	7	7	6
Black	89	86	5	5	5
Hispanic	25	24	3	2	3

¹The economically disadvantaged nonblack/non-Hispanic supplemental sample was dropped after the 1990 interview.

²Two hundred and one members of the military sample were

retained for continued interviewing after the sample was dropped in 1985; 183 of these respondents were interviewed in 1990.

Table 3.2. Number of NLSY79 civilian respondents by type of multiple respondent household (1979)

Type of household	Number of respondents
Sibling	5,914
Two siblings	3,398
Three siblings	1,743
Four siblings	624
Five siblings	125
Six siblings	24
Spouse	334

Note: Siblings may be biological, step, or adopted. Some households may include both siblings and spouses, as well as respondents with other relationships not presented in this table.

The interviews

Interviews with NLSY79 respondents were conducted on an annual basis from 1979 through 1994. After 1994, the survey schedule was changed to every other year. The original interview schedule, calling for in-person interviews, was maintained from 1979–86. In 1987, budget constraints dictated a more limited telephone interview; personal interviews resumed with the 1988 round. From 1979 to 1986, NLSY79 interviews were generally conducted in the first half of the year. Beginning in 1987, the fielding period shifted to the summer and fall months. Researchers should carefully note interview dates and fielding periods when using variables that may be affected by the time of year of the interview.

In 1993, computer-assisted personal interviewing (CAPI) replaced the traditional paper-and-pencil interview (PAPI) instruments used during the 1979–92 interviews. Experimental CAPI administrations had taken place during the 1989 and 1990 interviews, with documented improvements in data quality (see Baker and Bradburn, 1992; Bradburn, Frankel et al., 1992; Olsen, 1991; Speizer and Dougherty, 1991).

Table 3.3 presents the type of survey (personal or telephone), the interview mode (PAPI or CAPI), the number of interviewed respondents by sample type, and resulting retention rates for each NLSY79 round.

Attrition and sample drops

Retention rates for NLSY79 respondents considered eligible for interview remained close to 90 percent during the first 16 interview rounds and were around 85 percent for rounds 17 and 18. Retention rate, expressed as a percentage, is calculated by dividing the number of respondents interviewed by the number of respondents remaining eligible for interview. All base-year respondents, including those reported as deceased, are considered eligible for interview;

those who have been permanently dropped from the sample are not considered eligible for interview. In the 1998 survey, 8,399 civilian and military respondents out of the 9,964 eligible were interviewed, for an overall retention rate of 84.3 percent. Retention rates for each survey year are shown in table 3.3.

The number of respondents interviewed can also be expressed as a percentage of the number of base-year respondents not known to be deceased. This number, called response rate, is reported in table 3.4 for each survey. As of 1998, 275 respondents (almost 3 percent of the respondents eligible for interview) had been reported as deceased. Thus, the response rate for those still believed to be alive is almost 87 percent.

Funding constraints required that selected respondents in the military and supplemental subsamples be dropped from interviewing following the 1984 and 1990 surveys, respectively. After the 1984 survey, interviewing ceased for 1,079 members of the military subsample; retained for continued interviewing were 201 respondents randomly selected from the entire military sample. After the 1990 survey, all 1,643 economically disadvantaged nonblack/non-Hispanic respondents in the supplemental sample were dropped from the interview rolls.

Sampling weights

After the initial round, sampling weights were constructed to adjust for differential nonresponse and for the black, Hispanic, and economically disadvantaged oversamples. These weights provide the researcher with an estimate of the number of individuals in the United States who are represented by each respondent's answers. NORC recalculates the weights after each survey to account for noninterviews.

Major data elements

Each round of the NLSY79 contains core sets of questions on the following topics: Work and nonwork experiences, training investments, schooling, family income and assets, health conditions and injuries, geographic residence and environmental characteristics, household composition, and marital and fertility histories. Although the surveys have collected information on these topics on a regular basis, users should note that the number of questions on a given topic, as well as the wording and universe(s) for each question, may differ from year to year.

Selected survey years have included additional sets of questions on a variety of factors potentially affecting labor force attachment. The initial survey, for example, collected information on each respondent's family background, his or her knowledge of the world of work, and a detailed retrospective evaluation of the respondent's labor market experience. Subsequent surveys have included questions on subjects such as job search methods, migration, school dis-

Table 3.3. NLSY79 interview schedules and retention rates by sample types

Year	Type and mode of interview ¹	Cross-sectional sample total/retention rate		Supplemental sample total/retention rate		Military sample total/retention rate		Total sample total/retention rate	
1979	Personal/PAPI	6,111	—	5,295	—	1,280	—	12,686	—
1980	Personal/PAPI	5,873	96.1	5,075	95.9	1,193	93.2	12,141	95.7
1981	Personal/PAPI	5,892	96.4	5,108	96.5	1,195	93.4	12,195	96.1
1982	Personal/PAPI	5,876	96.2	5,036	95.1	1,211	94.6	12,123	95.6
1983	Personal/PAPI	5,902	96.6	5,093	96.2	1,226	95.8	12,221	96.3
1984	Personal/PAPI	5,814	95.1	5,040	95.2	1,215	94.9	12,069	95.1
1985	Personal/PAPI	5,751	94.1	4,957	93.6	² 186	92.5	² 10,894	93.9
1986	Personal/PAPI	5,633	92.2	4,839	91.4	183	91.1	10,655	91.8
1987	Telephone/PAPI	5,538	90.6	4,768	90.1	179	89.1	10,485	90.3
1988	Personal/PAPI	5,513	90.2	4,777	90.2	175	87.1	10,465	90.2
1989	Personal/PAPI/CAPI	5,571	91.2	4,853	91.7	181	90.0	10,605	91.4
1990	Personal/PAPI/CAPI	5,498	90.0	4,755	89.8	183	91.0	10,436	89.9
1991	Personal/PAPI	5,556	90.9	³ 3,281	89.9	181	90.0	³ 9,018	90.5
1992	Personal/PAPI	5,553	90.9	3,280	89.8	183	91.0	9,016	90.5
1993	Personal/CAPI	5,537	90.6	3,293	90.2	181	90.0	9,011	90.4
1994	Personal/CAPI	5,457	89.3	3,256	89.2	178	88.6	8,891	89.2
1996	Personal/CAPI	5,290	86.6	3,171	86.8	175	87.1	8,636	86.7
1998	Personal/CAPI	5,159	84.4	3,065	83.9	175	87.1	8,399	84.3

Note: Retention rate is defined as the percentage of base-year respondents remaining eligible who were interviewed in a given survey year; deceased respondents are included in the calculations.

¹PAPI interviews are those conducted with paper survey instruments and pencil-entered responses; CAPI interviews use a laptop computer and an electronic questionnaire that captures respondent-, interviewer-, and machine-generated data.

²A total of 201 military respondents were retained from the

original sample of 1,280; 186 of the 201 participated in the 1985 interview. The total number of NLSY79 civilian and military respondents eligible for interview beginning in 1985 was 11,607.

³The 1,643 economically disadvantaged nonblack/non-Hispanic male and female members of the supplemental subsample were not eligible for interview as of the 1991 survey year. The total number of NLSY79 civilian and military respondents eligible for interview beginning in 1991 was 9,964.

cipline, child care, self-esteem, time use, delinquency, knowledge about AIDS, attitudes toward work, childhood residences, neighborhood problems, drug and alcohol use, educational/occupational aspirations, and prenatal and post-natal health behaviors.

In addition to these core and supplemental data collections, a number of special surveys have focused on NLSY79 respondents. These studies include the 1980 high school survey, the 1980–83 high school transcript data collections, and the 1980 administration of the *Armed Services Vocational Aptitude Battery (ASVAB)*.

This section briefly describes major data elements for the NLSY79. The categories encompass information available not only from the NLSY79 main files but also from the NLSY79 work history and geocode constructed data files. This listing is by no means comprehensive, and not all data

elements are necessarily present for all respondents in all survey years. To determine whether the specific variables of interest are available for the year(s) and respondent universe(s) of interest, prospective users should examine several sources. They first should look at the detailed variable table in this chapter, which presents a more comprehensive look at the types of variables available on the NLSY79 main data files. Additional information can then be found in the NLSY79 survey instruments, the *NLSY79 User's Guide*, and the documentation files on the NLSY79 compact discs.

Data elements for the NLSY79 are discussed in the order presented in figure 3.1.

1. Work and nonwork experiences

Information is available for all respondents on their current

Table 3.4. NLSY79 response rates (excluding deceased respondents) by sample types

Year	Cross-sectional sample			Supplemental sample			Military sample			Total sample		
	Total	# deceased	Response rate	Total	# deceased	Response rate	Total	# deceased	Response rate	Total	# deceased	Response rate
1979	6,111	0	—	5,295	0	—	1,280	0	—	12,686	0	—
1980	5,873	4	96.2	5,075	5	95.9	1,193	0	93.2	12,141	9	95.8
1981	5,892	15	96.7	5,108	14	96.7	1,195	0	93.4	12,195	29	96.3
1982	5,876	24	96.5	5,036	19	95.5	1,211	1	94.7	12,123	44	95.9
1983	5,902	27	97.0	5,093	26	96.7	1,226	4	96.1	12,221	57	96.8
1984	5,814	30	95.6	5,040	33	95.8	1,215	4	95.2	12,069	67	95.6
1985	5,751	36	94.7	4,957	43	94.4	¹ 186	0	92.5	¹ 10,894	79	94.5
1986	5,633	43	92.8	4,839	51	92.3	183	1	91.5	10,655	95	92.6
1987	5,538	51	91.4	4,768	56	91.0	179	3	90.4	10,485	110	91.2
1988	5,513	56	91.0	4,777	68	91.4	175	3	88.4	10,465	127	91.2
1989	5,571	60	92.1	4,853	78	93.0	181	3	91.4	10,605	141	92.5
1990	5,498	67	91.0	4,755	82	91.2	183	3	92.4	10,436	152	91.1
1991	5,556	76	92.1	² 3,281	65	91.5	181	4	91.9	² 9,018	145	91.8
1992	5,553	81	92.1	3,280	71	91.6	182	4	92.4	9,015	156	91.9
1993	5,537	90	92.0	3,293	83	92.3	181	4	91.9	9,011	177	92.1
1994	5,457	104	90.8	3,256	96	91.6	178	4	90.4	8,891	204	91.1
1996	5,290	129	88.4	3,171	109	89.5	175	5	89.3	8,636	243	88.8
1998	5,159	152	86.6	3,065	118	86.7	175	5	89.3	8,399	275	86.7

Note: Response rate is defined as the percentage of base-year respondents remaining eligible and not known to be deceased who were interviewed in a given survey year.

¹A total of 201 military respondents were retained from the original sample of 1,280; 186 of the 201 participated in the 1985 interview. The total number of NLSY79 civilian and military respondents eligible for interview (including deceased respondents) beginning in 1985 was 11,607.

deceased) beginning in 1985 was 11,607.

²The 1,643 economically disadvantaged nonblack/non-Hispanic male and female members of the supplemental subsample were not eligible for interview as of the 1991 survey year. The total number of NLSY79 civilian and military respondents eligible for interview (including deceased respondents) beginning in 1991 was 9,964.

labor force status (that is, whether the respondent was employed, unemployed, or out of the labor force during most of the week preceding the survey). Those who report being employed supply job- and company-specific data, such as occupation, industry, firm size, availability of fringe benefits, job satisfaction, and number of hours worked each week. Detailed job information also is available on up to five employers for whom the respondent worked since the last interview. The surveys ask about the dates on which the job started and stopped, the number of hours worked, the reason why the respondent left the job, details on up to four gaps in employment while associated with an employer, occupation, class of worker, rate of pay, and collective bargaining activity in setting wages. Beginning with the 1994 survey, the information collected about each employer was expanded to include the job search method(s) that the re-

Figure 3.1. Data elements in the NLSY79

1	Work and nonwork experiences
2	Training investments
3	Schooling, school records, and aptitude information
4	Military experience
5	Income and assets
6	Health conditions, injuries, and insurance coverage
7	Alcohol and substance use, criminal behavior
8	Attitudes and aspirations
9	Geographic residence information
10	Family background and demographic characteristics
11	Household composition
12	Marital and fertility histories
13	Child care

spondent used to obtain the job; the types of fringe benefits provided by each employer; whether the respondent participated in a pension plan available through that employer, and whether the respondent received severance pay when leaving the job. The 1994–98 surveys also asked whether the position held with the employer was permanent, temporary, or as an independent contractor.

Also available is information on the activity of the respondent during periods when he or she was not associated with an employer or was in the military (that is, weeks not working, weeks spent looking for work, reasons not looking for work). Data are collected on job search behavior for unemployed respondents and on plans to seek employment for those out of the labor force. The surveys also have gathered information on the type and success of various job search methods used by unemployed respondents to find work. The 1979 interview included hypothetical job offer and reservation wage information.

CHRR has constructed a week-by-week longitudinal work record of each respondent from January 1, 1978, through the most recent survey date; this record is available as a separate data file (NLSY79 Work History CD). These weekly data are arranged in three primary arrays:

- A STATUS array of the respondent's labor force/military activity each week, beginning in January 1978
- An HOUR array of the usual hours worked per week at all jobs, beginning in January 1978
- A DUALJOB array, containing up to four additional job numbers for respondents who worked at more than one job simultaneously in any week, beginning in January 1978

Data on the work history disc include information about the number of weeks of a respondent's active military service, the number of weeks that the respondent spent in various labor force states, and detailed information regarding up to five civilian jobs reported by the respondent during each survey year. The respondent identification number permits researchers to link these data with the main youth, geocode, and child files.

2. Training investments

The surveys have regularly collected detailed information about the types of nongovernment-sponsored vocational or technical training programs in which a respondent had enrolled since the last interview. In general, the scope of information gathered includes the occupation in which a respondent received training, dates of enrollment and completion, the type of school providing the training, whether the respondent received any subsequent training, the types of certificates or licenses he or she earned, and how the respondent paid for the training.

The 1979–86 surveys asked an extensive series of questions about government-sponsored training programs in which the respondent had enrolled since the last interview. This series includes program-specific information regarding the respondent's current enrollment status, dates/hours of participation, periods of nonparticipation, and whether the program was part of a federally funded training program such as Job Training Partnership Act (JTPA)/CETA or the Work Incentive Program (WIN). These surveys also collected information on the type of occupational or on-the-job training received, the types of classroom training and supportive services provided, and the respondent's rate of pay during training. For jobs identified as government-sponsored, information includes the job's occupational code, the types of classroom training received, and supportive services provided in conjunction with the job and job placement information. Surveys conducted after 1986 posed substantially fewer questions about respondents' participation in government training and employment programs.

In 1993 and 1994, the interviews included a series of questions on (1) the informal methods (for instance, classes/seminars, supervisors, coworkers, self-study) used to learn skills required on the respondent's current job and (2) the potential transferability of skills acquired during various on- and off-the-job training programs. The 1993 survey also fielded a set of high school course relevancy questions. This series included questions on whether the skills learned in a training program added to those the respondent had acquired during high school. For those who did not take such high school courses, questions attempted to determine how much of what was learned in the training program could have been learned in high school. Beginning in 1996, questions on informal training on current jobs in the last 4 weeks replaced the questions about informal training used to learn the job.

3. Schooling, school records, and aptitude information

NLSY79 respondents regularly answer questions about their schooling investments since date of last interview. Information available includes respondents' current school enrollment status, the highest grade they have attended or completed, whether they have earned a high school diploma or GED equivalency, the type of their high school curriculum, their college enrollment status, their major field of study at college, and the type(s) of college degrees they have earned. The 1994 survey collected retrospective information about respondents' participation in Head Start or another preschool program.

The 1980 high school survey gathered data from the records of the last secondary school attended by NLSY79 respondents; data are available for more than 8,500 respondents. This set of variables includes both respondent-specific and school-specific information about factors such as (1) the respondent's school enrollment status, highest grade

attended, remedial classes taken, and scores for various intelligence and aptitude tests that had been administered during the youth's schooling and (2) each school's total enrollment, grading system, types of curricula offered, drop-out rate, student body composition, and staffing characteristics.

Transcript data collected during the 1980–83 survey years are available for more than 9,000 NLSY79 respondents who had completed or were expected to complete high school. Information about as many as 64 high school courses includes course titles, final grades, and credit received. The transcript data collection also provided respondents' attendance records, class rank, and scores on various aptitude and achievement tests.

In addition to the aptitude and intelligence scores collected during the high school and transcript surveys described above, scores from the *Armed Services Vocational Aptitude Battery (ASVAB)* are available for 11,914 NLSY79 respondents. These data include individual respondent raw scores, standard scores, scale scores, and standard errors for each of the 10 test sections (for example, general science, arithmetic reasoning, word knowledge, mechanical comprehension). CHRR created two Armed Forces Qualifications Test (AFQT) scores, which also are present in the data files.

4. Military experience

The survey provides information regarding the respondent's enlistment intentions, attitudes toward the military, dates of military service/reserve duty, branch of service, military occupation, pay grade, income, education/training received, and reasons for leaving the military or reenlisting. Surveys after 1985 use an abbreviated set of military experience questions.

5. Income and assets

The interviews record information about the sources and amounts of income received during the past calendar year by the respondent, his/her spouse, or other family members. Data on income sources of an opposite-sex partner are available beginning in 1990; if a respondent volunteers such information for a same-sex partner, that information also is recorded. Income sources identified include the respondents' and their spouses' or partners' wages and salaries, income from military service, profits from a farm or owned business, Social Security, pensions and annuities, and alimony/child support. The survey also asks about monthly income received by the respondent and spouse or partner from unemployment compensation, Aid to Families with Dependent Children (AFDC/TANF), Food Stamps, and other public assistance, including Supplemental Security Income. The specific sources of income received by other family members are not identified in all cases. Surveys since 1993 have asked a new set of child support questions that

allows comparisons to be made between amounts of support moneys to which a respondent or spouse was legally entitled and amounts of child support actually received. New public assistance questions added in 1998 asked respondents whether they had enrolled in training, looked for a job, or gotten a job to fulfill new welfare requirements; applied for public assistance since August 1996, when the new requirements went into effect; or been affected in some other way by the 1996 welfare reform. This series was shortened in 2000. A new question series in 2000 asked about receipt of targeted benefits, such as job or skills training, clothing assistance, or food assistance other than Food Stamps, WIC, or school meals.

In addition to the information described above, the data include a set of created variables detailing the respondent's history of participation in unemployment compensation (respondent and spouse), AFDC/TANF, SSI, and Food Stamp programs. For each type of reciprocity program, variables include: (1) A monthly dollar value of the benefits, (2) a yearly indicator of receipt or nonreceipt, and (3) a yearly dollar value of the benefits. In addition, there are two yearly summary variables that indicate whether the respondent received any form of public assistance during the year and the total dollar value of all benefits received from AFDC/TANF, SSI, and Food Stamp programs.

Asset and liability information collected during the 1985–90 and 1992–2000 survey years includes (1) the types and total market value of property owned by the respondent; (2) the value of other assets, including vehicles and savings accounts; and (3) the total amount of debts owed, including mortgages, back taxes, and other debts over \$500.

6. Health conditions, injuries, and insurance coverage

The surveys provide data regarding the respondent's height and weight, as well as the presence and duration of health conditions that prevent or limit labor market activity. Respondents also answer questions about type(s) of work-related injuries or illnesses they had experienced and their impact on the respondent's work activity. Information about each work-related injury includes the specific type of condition, causes, and part(s) of the body affected. The *International Classification of Diseases* coding system identifies health conditions and work-related injuries reported by respondents during selected survey years. In some years, mothers provide data on prenatal health care, infant feeding practices, infant illnesses and treatment, and well-baby care, as well as accidents and injuries to their children.

Recent surveys have collected information on whether NLSY79 respondents, their spouses, and their children are covered by a private or governmental health care/hospitalization plan and the source(s) of the coverage (current or previous employer, direct purchase, public assistance health care program, or other). The 1994 survey asked about spe-

cific periods without coverage for respondents, their spouses, and their children; surveys since 1994 have asked only about specific periods without coverage for respondents.

A set of health questions first asked in the 1998 survey establishes baseline health data for respondents aged 40 and older. Each subsequent survey addresses the health module to respondents who have reached age 40 but have not yet answered these questions. This extensive module records these respondents' general health status, the ways in which their health influences daily activities, and their emotional well-being in the past 4 weeks. Respondents also state whether they have experienced any of a long list of potential health problems, such as heart disease, cancer, diabetes, back pain, and so forth. Finally, limited data are collected on the health of the respondents' parents, including their life status, current state of health, and cause of death, if deceased.

7. Alcohol and substance use, criminal behavior

Alcohol use questions, asked during selected survey years, cover the respondent's consumption of alcohol, his or her frequency of use, quantity usually consumed, whether such use has affected the respondent's schoolwork or job performance, and whether any relatives have alcohol-related problems. Respondents also have provided information on their use of cigarettes, including the age at which they began smoking, how frequently they smoke, and the age at which they stopped smoking, if applicable. Substance use data include the respondent's age at first use; extent of use of marijuana/hashish, amphetamines, barbiturates, cocaine, and heroin; and the respondent's use of such substances on the job. Female NLSY79 respondents have regularly reported on their use of alcohol, cigarettes, marijuana, or cocaine during pregnancy.

A self-report supplement to the 1980 survey collected data on the respondent's participation in and income from various delinquent and criminal activities such as truancy, alcohol/marijuana use, vandalism, shoplifting, drug dealing, and robbery. A second set of questions in this supplement measured involvement with the criminal justice system, including reported arrest records and police contact.

8. Attitudes and aspirations

Attitude variables available for selected survey years include the *Internal-External Locus of Control Scale* (Rotter, 1966), the *Mastery Scale* (Pearlin et al., 1981), the *Self-Esteem Scale* (Rosenberg, 1965), and the *Center for Epidemiologic Studies Depression (CES-D) Scale* (Radloff, 1977). In some years, the survey has included questions on attitudes toward women and work; occupational aspirations; work commitment; knowledge of the world of work; perceived problems in getting a good job; and future expectations about mar-

riage, education, and employment. In the initial survey, respondents reported on the attitude of the most influential person in their life toward certain key career, occupational, residence, and childbearing decisions.

9. Geographic residence information

The main NLSY79 data files contain general information on the larger environment in which the respondent lives and works. This information includes details on each respondent's country, State, and region of residence at birth and at age 14; region of residence at each interview date; whether the current residence is urban or rural; and whether the residence is located in a Standard Metropolitan Statistical Area (SMSA). The 1992 and 1994–2000 surveys collected neighborhood characteristics information from NLSY79 women, asking them to rate their neighborhood as a place to raise children and to identify the extent of neighborhood problems such as crime and violence, lack of police protection, neighborhood apathy, and joblessness.

The restricted-release, detailed geocode data files provide additional geographic information for NLSY79 respondents. These separate data files are distributed only to NLS researchers who satisfactorily complete the BLS geocode accessing agreement procedure. The geocode file includes information on the respondent's State, county, and SMSA/MSA/CMSA/PMSA of current residence; the location of the most recent college that the respondent attended; and selected environmental variables from the *County and City Data Books* for the respondent's county or SMSA of current residence. For information about receiving a geocode disc, researchers should contact BLS, 2 Massachusetts Ave. NE, Room 4945, Washington, DC 20212-0001; (202) 691-7410; e-mail: NLS_Info@bls.gov.

A separate data set, the NLSY79 Women's Support Network data file, offers measures of the geographic proximity of the residences of relatives, friends, and acquaintances to female NLSY79 respondents interviewed during the 1983 through 1985 surveys.

10. Family background and demographic characteristics

Selected surveys have collected information on respondents' gender (interviewer-coded), racial/ethnic identification, date of birth, State or country of birth, number of siblings, religious affiliation, childhood residences from birth to age 18, and immigration/visa status, as well as background information about their parents' birthplace, education, and work experiences. The 1993 survey asked respondents to report the age, education, and fertility of as many as 13 biological siblings. For those respondents identifying a sibling of the same age, follow-up questions in 1994 confirmed whether the paired relationship was that of an identical or fraternal twin, triplet, or quadruplet.

11. Household composition

Basic demographic information is available for each person living in the respondent's household at the time of the survey. These data include each resident's sex, relationship to the respondent, age, highest grade of schooling completed, and work experience during the past year.

12. Marital and fertility histories

Regularly collected information includes respondents' marital status at the survey date, changes in marital status since the last interview, and dates of each marital status change. Certain surveys have gathered information about the respondents' spouses (and, on a more limited basis, partners), including birth and death dates, occupation, educational attainment, labor force status, religious affiliation, and health limitations. Since 1994, "partner" is generally considered equal to "spouse."

Fertility data collected for NLSY79 women include information about all pregnancies (through 1991) or live births (1992–present), a cumulative roster of all children, residence status of all children, contraceptive methods used, birth expectations and wantedness information, confidential abortion reports, and ages at menarche and first sexual intercourse. Fertility data collected for men in the NLSY79 cohort are more limited. Information includes birth data for the respondent's children, the residence status of his children, his age at first sexual intercourse, and contraceptive methods used by the respondent. Additionally, a set of questions asked in 1998 and 2000 recorded the frequency of the male respondent's various interactions with his children.

A set of edited fertility variables added to the main data file (area of interest FERTILE) provides revisions to dates of birth, gender, and usual living arrangements for all respondents' children, as well as flag variables indicating the types of revisions made. A set of constructed variables includes the beginning and ending dates of marriages, ages at first marriage and first birth, spacing between births, and spacing between marriage and first birth for the respondent's children. Because they have been corrected, variables in this area of interest are more accurate than the raw data. Users are strongly encouraged to use these edited and created variables. Finally, the database contains a variable evaluating the consistency of each respondent's longitudinal fertility record between the 1979 and 1982 survey years.

13. Child care

Selected surveys provide information about the types of child care used by female respondents (for instance, care by relatives, care by nonrelatives, day care center, nursery or preschool, child cares for self), the types of child-care payments incurred, and the number of hours or days that the respondents' children spent in child care. Retrospective data

collected in recent surveys include child-care experiences and arrangements in the first 3 years of life for all children at least 1 year old.

Data files and documentation

Data for the NLSY79 from 1979–98 are distributed on three sets of compact discs: (1) An NLSY79 Main File 1979–1998 two-disc set, which excludes the restricted geocode data; (2) an NLSY79 Geocode File 1979–1998 two-disc set with all main file data and geocode data; and (3) an NLSY79 Work History Disc 1979–1998. The contents of each disc are described in chapter 8 of this handbook.

Each disc includes documentation files and search and extraction software that enables users to peruse, select, and extract variables easily. NLS compact discs are designed for use in a PC MS-DOS environment. Chapter 8 contains more information on the hardware requirements of the NLS compact discs and descriptions of supplemental NLSY79 documentation items not present on the disc.

References

- Baker, Reginald P. and Bradburn, Norman M. "CAPI: Impacts on Data Quality and Survey Costs." In: *Proceedings of the 1991 Public Health Conference on Records and Statistics*. Washington, DC: U.S. National Center for Health Statistics (1992): 459–64.
- Bradburn, Norman M.; Frankel, Martin R.; Baker, Reginald P.; and Pergamit, Michael R. "A Comparison of Computer-Assisted Personal Interviews (CAPI) with Paper-and-Pencil (PAPI) Interviews in the National Longitudinal Study of Youth." *NLS Discussion Paper* 92-2. Washington, DC: U.S. Department of Labor, Bureau of Labor Statistics, 1992.
- Olsen, Randall J. "The Effects of Computer Assisted Interviewing on Data Quality." Columbus, Ohio: CHRR, The Ohio State University, 1991.
- Pearlin, Leonard I.; Lieberman, Morton A.; Menaghan, Elizabeth G.; and Mullan, Joseph T. "The Stress Process." *Journal of Health and Social Behavior* 22 (December 1981): 337–56.
- Radloff, Lenore Sawyer. "The CES-D Scale: A Self-Report Depression Scale for Research in the General Population." *Applied Psychological Measurement* 1,3 (Summer 1977): 385–401.
- Rosenberg, Morris. *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press, 1965.
- Rotter, Julian B. "Generalized Expectancies for Internal

vs. External Control of Reinforcements.” *Psychological Monographs* 80,1 (1966): 1–28.

Speizer, Howard and Dougherty, Doug. “Automating Data Transmission and Case Management Functions for a Nationwide CAPI Study.” In: *Proceedings of the 1991 Annual Research Conference*. Washington, DC: Census Bureau (1991): 389–97.

Selected variables by survey year: NLSY79

Table 3.5 depicts selected variables from the NLSY79 main data files. Additional raw and constructed variables on a number of topics (such as geographic-environmental data and detailed individual work histories) can be found on the

separately released NLSY79 geocode disc and on the work history disc. Detailed information on the Children of the NLSY79 can be found in the various handbooks, user’s guides, and evaluation reports that document the child data; more information is available in chapter 4.

The top row of the table lists the years in which the NLSY79 has been administered; the first column gives some of the variables for which information has been collected. An asterisk (*) indicates that information for the designated variable was obtained for the given cohort in the year indicated. “R” used in variable descriptions stands for “respondent.” Users should be aware that, in some instances, the variable descriptions in the table represent a group of questions and not single response items.

Table 3.5. Selected NLSY79 variables by survey year: Respondents aged 14 to 22 in 1979

Variable	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	96	98	00
I. LABOR MARKET EXPERIENCE VARIABLES																			
A. Current labor force and employment status and characteristics of current/most recent job¹																			
Survey week labor force and employment status	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Occupation (DOT code)	*																		
Hours worked in survey week	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Hours per week usually worked	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Shift worked	*	*	*	*	*	*	*			*	*	*	*	*	*	*	*	*	*
Promotion (varies with year)						*	*	*	*	*	*	*					*	*	*
Commuting time to current job	*	*	*							*					*	*			
Availability of fringe benefits	*	*		*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Global job satisfaction item	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Job satisfaction scale	*	*	*	*						*									
Job characteristics inventory	*			*															
Size of employer	*	*						*	*	*	*	*	*	*	*	*	*	*	*
Minority status of coworkers (1980 only), supervisor		*		*															
Time R expects to stay at job	*	*	*	*															
Participation in work-study program	*	*	*	*	*	*	*	*											
B. Work experience since Jan. 1, 1978, previous survey, or in past calendar year																			
Weeks worked	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Hours usually worked per week	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Number of weeks, spells of unemployment	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Weeks out of labor force	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
C. Characteristics of jobs since Jan. 1, 1978, or last survey, including current or most recent job (more than 10 to 20 hours per week and more than 9 weeks in duration if not CPS job)																			
Occupation and industry (census code)	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Class of worker	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Start date and stop date	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Hours usually worked at home										*	*	*	*	*	*	*	*	*	*
Shift worked																*	*	*	*
Rate of pay, hourly rate of pay	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Covered by collective bargaining	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Is R union member	*									*	*	*	*	*	*	*	*	*	*
Reason for leaving job	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Severance pay received																*	*	*	*
Availability of fringe benefits																*	*	*	*
Characteristics of employer's pension plan																*	*	*	*

¹ Beginning in 1994, characteristics of the current or most recent job were collected in the first Employer Supplement loop, rather than in the CPS section. To maintain consistency, these questions are still included in this section of the table.

Table 3.5. Selected NLSY79 variables by survey year: Respondents aged 14 to 22 in 1979

[illegible]

Table 3.5. Selected NLSY79 variables by survey year: Respondents aged 14 to 22 in 1979

Variable	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	96	98	00
High school curriculum	*	*	*	*	*	*	*												
Comparison of high school courses to skills training															*				
College degree received	*	*	*	*	*	*				*	*	*	*	*	*	*	*	*	*
Type of college attending (2- or 4-year)	*	*	*	*	*	*	*	*		*	*	*		*	*	*	*	*	*
Field of study or specialization in college	*	*	*	*	*	*	*	*		*	*	*		*	*	*	*	*	*
College tuition	*																		
Educational loans or financial aid in college	*	*	*	*	*	*	*	*		*	*	*		*	*	*	*	*	*
Attitude toward selected aspects of high school	*																		
Courses taken during last year of high school	*																		
Ever suspended or expelled from school; date		*																	
C. Vocational training outside regular school																			
Type(s) of training	*	*	*	*	*	*	*	*		*	*	*	*	*	*	*	*	*	*
Number of weeks, hours per week in training	*	*	*	*	*	*	*	*		*	*	*	*	*	*	*	*	*	*
Was training completed	*	*	*	*	*	*	*	*		*	*	*	*	*	*	*	*	*	*
Was degree, certificate, or journeyman's card obtained	*	*																	
Was training related to specific job or employer				*	*	*	*	*		*	*	*	*	*	*	*	*	*	*
Was training related to a promotion												*	*	*	*	*			
Reason for training				*	*	*						*	*	*	*	*	*	*	*
Method of financing training				*	*	*				*	*	*	*	*	*	*	*	*	*
Informal job learning activities															*	*	*	*	*
D. Government jobs and training programs																			
Participation in programs	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Type of program	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Satisfaction with program	*	*	*	*	*	*	*	*	*										
Did program help on subsequent jobs	*	*	*	*	*	*	*	*	*		*	*	*	*	*	*	*	*	*
Services provided by program	*	*	*	*	*	*	*	*	*										
Length of participation in program	*	*	*	*	*	*	*	*		*	*	*	*	*	*	*	*	*	*
Hours per week and per day spent in program	*	*	*	*	*	*	*	*		*	*	*	*	*	*	*	*	*	*
Amount of income from participating in program	*	*	*	*	*	*	*	*											
Aspects liked most and least about programs	*																		
Reasons for entering and leaving programs	*	*	*	*	*	*	*	*	*										
E. Health and physical condition																			
Does health limit work, duration of limitation	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Type of health problem (ICD-9 code)	*	*	*	*															
Work-related injury or illness (ICD-9 code)										*	*	*		*	*	*	*	*	*
Height			*	*			*												
Weight			*	*			*	*		*	*	*	*	*	*	*	*	*	*
Health insurance coverage—R, spouse, children											*	*		*	*	*	*	*	*

Table 3.5. Selected NLSY79 variables by survey year: Respondents aged 14 to 22 in 1979

Variable	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	96	98	00
Frequency and intensity of R's physical activity																		*	*
General perception of health (40-and-older health module)																		*	*
Does health interfere with daily activities (40-and-older module)																		*	*
Emotional health in past 4 weeks (40-and-older module)																		*	*
R's various health problems (heart problems, cancer, diabetes, poor eyesight or hearing, and so forth) (40-and-older module)																		*	*
Time spent on health care activities (40-and-older module)																		*	*
F. Marital history and spouse characteristics																			
Dating behaviors and attitudes (unmarried females)										*				*		*	*	*	*
Marital status	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Changes in marital status since 1/1/1978 or previous interview; number and duration of marriages	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Month, year R and partner began living together												*	*	*	*	*	*	*	*
Did R and spouse live together continuously before marriage (or R and partner continuously until now)												*	*	*	*	*	*	*	*
Occupation of spouse (or partner, since 1994)	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Extent spouse worked in previous calendar year (or partner, since 1994)	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Current labor force status, reason not employed for spouse (or partner, since 1994)										*	*	*	*	*	*	*	*	*	*
Shift worked by spouse (or partner, since 1994)				*						*	*	*	*	*	*	*	*	*	*
Rate of pay, hourly rate of pay of spouse (or partner, since 1994)												*	*	*	*	*	*	*	*
Spouse's religious affiliation and attendance (or partner in 2000)				*															*
Number of spouse's marriages, details (or partner, in 1998 and 2000)				*														*	*
Effect of spouse's health on R's work				*															
Quality of R's relationship (14 items)										*				*		*	*	*	*
Age at which R expects to marry	*																		
G. Household and children																			
Relationship of household or family members to R	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Household or family members' demographics (sex, age, highest grade completed, work status in past year)	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Number of dependents or exemptions	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Number and ages of children living in household	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Expected number of children	*			*	*	*	*	*		*		*		*		*	*	*	*

Table 3.5. Selected NLSY79 variables by survey year: Respondents aged 14 to 22 in 1979

Variable	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	96	98	00
Number of children R considers ideal	*			*															
Health care during pregnancy (females)					*	*	*	*		*		*		*		*	*	*	*
Postnatal infant care and feeding (females)					*	*	*	*		*		*		*		*	*	*	*
Medical care received by children (females)					*	*	*	*		*		*		*		*	*	*	*
Father's relationship with children (males)																		*	*
Fertility history	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Use of birth control methods				*		*	*	*		*		*		*		*	*	*	*
Pregnancies not resulting in live births (includes how ended through 1990)				*	*	*	*	*		*		*		*		*	*	*	*
H. Child care (only female R's in even years after 1986)																			
Current child-care arrangements				*	*	*	*	*	*	*									
Child care during first 3 years of life								*		*				*		*	*	*	*
Cost per week				*			*	*		*									
Number of hours per week				*	*	*	*	*		*									
Is child care a hindrance to R's work, school, or training				*	*	*				*	*								
Extent of various neighborhood problems														*		*	*	*	*
I. Financial characteristics and program participation																			
Total family income in previous calendar year	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Income of R and spouse in previous calendar year from:																			
farm or own business	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
wages or salary	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Unemployment compensation	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
public assistance	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Food Stamps	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
targeted cash or noncash benefits																		*	
pensions/Social Security	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
military service	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
veteran's benefits, workers' compensation, other disability	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
other sources	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
R receives government rent subsidy or public housing	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Income from child support	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Child support expected vs. received															*	*	*	*	*
Possession of various assets (R and spouse)		*	*	*	*	*	*	*	*	*	*	*		*	*	*	*	*	*
Asset market value (R and spouse)							*	*	*	*	*	*		*	*	*	*	*	*
Amount of debt							*	*	*	*	*	*		*	*	*	*	*	*
Amount spent on food, other than Food Stamps												*	*	*	*	*			

Table 3.5. Selected NLSY79 variables by survey year: Respondents aged 14 to 22 in 1979

[illegible]

Table 3.5. Selected NLSY79 variables by survey year: Respondents aged 14 to 22 in 1979

Variable	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	96	98	00
CES-Depression Scale (R's aged 40 and older in 1998, 2000)														*		*		*	*
Person having most influence on R, his or her responses to various situations	*																		
M. Retrospective evaluation of labor market experience																			
Perception of age, race, and sex discrimination	*			*															
Reason for problems in obtaining employment	*			*															
N. Delinquency, drugs, and alcohol use																			
Activities within last year (20 items)		*																	
Income from illegal activities within last year		*																	
Alcohol consumption in last week or month				*	*	*	*			*	*			*		*			
Extent of cigarette use						*								*		*		*	
Age R first smoked and stopped smoking cigarettes														*		*		*	
Extent of marijuana use		*				*				*				*		*		*	
Age R first used marijuana						*				*				*		*		*	
Extent of cocaine use, age R first used						*				*				*		*		*	
Extent of "crack" cocaine use, age R first used														*		*		*	
Ever used sedatives, barbiturates, and so forth						*								*		*		*	
Cigarette and alcohol use during pregnancy					*	*	*	*		*		*		*		*	*	*	*
Marijuana and cocaine use during pregnancy										*		*		*		*	*	*	*
O. Reported police contacts																			
Number of times stopped by police		*																	
Number of times booked or arrested		*																	
Number of convictions, charges		*																	
Number of times incarcerated; date of release		*																	
P. Time use																			
Use of time at various activities (school, work, watching TV, household chores, and so forth)			*																
III. ENVIRONMENTAL VARIABLES																			
Place of birth (South or nonsouth)	*																		
Residence at age 14 (South or nonsouth, and State or country)	*																		
Region of residence	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Current residence urban or rural	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Current residence in SMSA	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Unemployment rate of local labor market	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Changes in residence since January 1, 1978, or date of last interview (collected as a history)	*	*		*															*

Note: Excluded from this table are the detailed geographic, fertility, and child-specific variables that appear on separate NLSY79 constructed data files.